DOCUMENT RESUME

ED 422 087 PS 026 744

TITLE	Head Start: Challenges in Monitoring Program Quality and Demonstrating Results. Report to Congressional Requesters.
INSTITUTION	General Accounting Office, Washington, DC. Health, Education, and Human Services Div.
REPORT NO	GAO/HEHS-98-186
PUB DATE	1998-06-00
NOTE	45p.
AVAILABLE FROM	U.S. General Accounting Office, P.O. Box 37050, Washington, DC 20013; phone: 202-512-6000; fax: 202-512-6061; e-mail: info@www.gao.gov; World Wide Web: http://www.gao.gov (First copy free. Additional copies, $2 each).
PUB TYPE	Legal/Legislative/Regulatory Materials (090) -- Reports - Evaluative (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Accountability; Comparative Analysis; Compliance (Legal); *Early Childhood Education; Educational Quality; *Evaluation Methods; Evaluation Research; Federal Programs; Preschool Education; *Program Evaluation; Research Problems
IDENTIFIERS	Department of Health and Human Services; Outcome Oriented Evaluation; Performance Indicators; Performance Monitoring; *Project Head Start; Quality Assurance

ABSTRACT

There has been increasing concern about the accountability of federal programs in complying with laws and regulations and in achieving desired results. At the request of Congress, the Government Accounting Office examined how the Department of Health and Human Services (HHS) ensures Head Start accountability. Data were collected by means of interviews with federal regional, and selected local officials; and reviews of selected grantees' program files. Findings indicated that Head Start's mission, goal, and objectives do provide an overall performance assessment framework emphasizing compliance with applicable laws and regulations and achievement of program results, with specific program activities linked to the strategic mission and goal. By specifying measurable performance objectives, Head Start can answer questions about compliance and achievement of its purpose. Although HHS has processes to ensure compliance, implementation could be improved. Of concern is the consistency of HHS on-site inspections of local agencies, due to differences in reviewers assessments. Although HHS has strengthened emphasis on achieving program purposes, the processes provide too little information about how well programs are achieving their intended purpose. New evaluation initiatives will provide information on program outcomes not previously available, such as gains in child vocabulary, literacy, and social skills and families' self-sufficiency. Although HHS will collect data from all programs, there is no established plan or schedule for doing so. The HHS planned analysis will not enable the Department to determine with certainty that Head Start participation caused improvements because results will not be compared to a group of children and families similar in all respects except Head Start participation. (KB)

United States General Accounting Office

GAO

Report to Congressional Requesters

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

June 1998

ED 422 087

HEAD START

Challenges in Monitoring Program Quality and Demonstrating Results

PS 026744

GAO/HEHS-98-186

2

GAO

United States
General Accounting Office
Washington, D.C. 20548

Health, Education, and
Human Services Division

B-278560

June 30, 1998

Congressional Requesters

Head Start, one of the most popular federal early childhood programs, has provided funding for a comprehensive set of services to about 16 million low-income preschool children in the past 33 years. Administered by the Department of Health and Human Services' (HHS) Administration for Children and Families (ACF), the program has long enjoyed both congressional and public support. Since its inception, Head Start costs have totaled $35 billion, and its annual funding has increased substantially in recent years. Between fiscal years 1990 and 1998, annual Head Start funding nearly tripled from $1.5 billion to almost $4.4 billion. The administration recently proposed a significant funding increase for Head Start to expand the program's annual enrollment to one million children by 2002. Meanwhile, the Congress, executive branch, and taxpayers have become more concerned about ensuring the accountability of federal programs in not only complying with laws and regulations, but also achieving desired results.

Given the significant financial investment in Head Start as well as its increased funding in recent years and the interest in holding federal programs accountable for their performance, you asked us to evaluate how HHS ensures that Head Start programs are accountable for complying with laws and regulations and for achieving program purposes. "Achieving program purposes" refers to whether desired outcomes have been achieved and whether Head Start participation has caused differences in outcomes. Specifically, we addressed

- the extent to which Head Start's mission, goal, and objectives provide an overall framework that emphasizes compliance with applicable laws and regulations and achievement of program results;
- how well Head Start's processes ensure compliance with applicable laws and regulations; and
- how well Head Start's processes ensure the ability to determine whether program purposes have been achieved.

We analyzed the Head Start program's mission statement, strategic goal, program objectives, and regulations. In addition, we examined monitoring processes, systems, and tools used to ensure accountability at (1) national headquarters, (2) regional offices that administer regular Head Start programs at the local level, and (3) regular Head Start local service

providers.[1] We interviewed federal headquarters officials who oversee and administer the Head Start program, including those in HHS' ACF and its Head Start Bureau, Division of Grants Policy, Division of Payment Management, and Office of Inspector General. We also visited 6 of the 10 ACF regional offices to (1) interview regional officials who administer and oversee the Head Start program and, (2) at 5 regional offices, review selected program files of grantees being monitored because of poor performance. In addition, we interviewed selected local Head Start program directors and staff to identify tools, processes, and systems used to ensure accountability in their local programs and monitor performance. We did our work between February 1997 and May 1998 in accordance with generally accepted government auditing standards.

Results in Brief

Head Start's mission, goal, and objectives provide an overall performance assessment framework that emphasizes compliance with applicable laws and regulations and achievement of program results. Head Start developed this framework, which reinforces the program's accountability by linking specific program activities to its overall strategic mission and goal, in response to legislative requirements, such as the Results Act, as well as Head Start Bureau policies. By specifying measurable program performance objectives, Head Start has the ability to answer questions about its compliance with regulations and whether it is achieving its purposes. This could help to answer the critical question of whether the program is having an impact, that is, making a difference in participants' lives.

Although HHS has processes in place to ensure that grantees comply with regulations, the implementation of these processes could be improved, according to our review. For example, HHS periodically conducts comprehensive on-site inspections of local Head Start agencies using a process that, if properly implemented, can ensure regulatory compliance. Both HHS' and our reviews, however, have identified concerns about the consistency of these inspections due to differences in reviewers' assessments of whether grantees are complying with some requirements and due to other factors. In essence, different reviewers may interpret inspection results differently. These different conclusions, in turn, may affect the decisions made about actions to take on the basis of inspection results.

[1]The regular Head Start program serves children and families in the 50 states, the District of Columbia, Puerto Rico, and the trust territories. About 85 percent of Head Start children are served through the regular Head Start program. Head Start also operates programs for migrant and Native American populations. Grantees and their delegate agencies are the local service providers.

In recent years, HHS has substantially strengthened its emphasis on determining whether it has achieved program purposes. Its processes provide too little information, however, about how well the program is achieving its intended purposes. HHS has new initiatives that, in the next few years, will provide information not previously available on program outcomes, such as gains made by participating children and their families. This information—from a survey of a nationally representative sample of families—will describe, for example, the extent to which participating children have improved their vocabulary, literacy, and social skills as well as the extent to which families have become economically and socially self-sufficient. In the future HHS will collect such data from all Head Start programs rather than from just a sample of programs, according to agency officials, but it has no established plan or schedule for doing so. Until HHS takes this step, it will continue to hold local Head Start programs accountable only for complying with regulations—not for demonstrating progress in achieving program purposes. Moreover, although HHS' survey will allow Head Start to show whether children and their families have progressed in achieving program purposes, HHS' planned analysis of survey results will not allow it to determine with certainty that Head Start participation caused children's or their families' improvements. Instead of comparing survey results with those from a group of children and families similar in all respects except for their Head Start participation, HHS will compare results with other groups. This approach will not allow HHS to isolate Head Start participation as a causal factor in children's and families' progress. Therefore, HHS will not be able to determine program impact, that is, whether the program is making a difference in children's and families' lives.

Background

Head Start was created in 1965 as part of President Johnson's War on Poverty. The program provides comprehensive services, such as educational, medical, nutritional, mental health, dental, and social services, to low-income children and their families in all 50 states, the District of Columbia, Puerto Rico, and the U.S. territories. Head Start also provides services for migrant and Native American populations. Unlike some other federal social service programs that are funded through the states, HHS awards Head Start grants directly to local agencies, called grantees, which deliver program services. Grantees numbered about 1,456 in fiscal year 1997. Grantees may contract with other organizations, called delegate agencies, to run all or a part of their programs. Grantees had such contracts with about 517 delegate agencies in fiscal year 1997; four grantees had contracts with 175 of these 517 delegate agencies. Grantees

and delegate agencies include public and private school systems, community action agencies and other private nonprofit organizations, local governments, and Indian tribes.

HHS distributes funds to grantees on the basis of, among other things, the amount of funds a particular grantee received in preceding years and a proportion of the amount of additional Head Start funds available. About 60 percent of all regular Head Start grantees have participated in the program for over 25 years. Once approved for funding through a competitive application process, grantees no longer compete for funding in future years. Grantees must, however, submit applications to continue receiving funds for their programs. Grantees must generally pay at least 20 percent of total program costs with nonfederal funds. State or local governments or private sources often provide such funds in the form of cash or in-kind contributions such as building space or equipment.

ACF administers the Head Start program through the Head Start Bureau and ACF's regional offices nationwide.[2] The Head Start Bureau develops program policy, goals, and objectives for the program and compiles reports on the program for the Congress and the public. ACF's regional offices implement the Head Start Bureau's policies as well as administer and oversee local Head Start agencies, which includes assessing local agencies' compliance with applicable laws and regulations.

Since its inception, Head Start has had long-standing and widespread support. In recent years, however, HHS' Inspector General and we have expressed concerns about program operations and effectiveness, and the Congress has acted to strengthen program accountability. For example, the Inspector General issued a series of reports in the early 1990s seeking to improve, among other things, grantees' financial management and HHS' program compliance capabilities. In addition, we have issued four reports in as many years on Head Start. One report, which addressed grantees' views on barriers to providing services, found an insufficient number of qualified grantee staff to meet the complex needs of children and their families.[3] Grantees were also having trouble finding suitable facilities they could afford, according to our review. Our second report found that Head Start centers often provide a broader range of services to children and

[2]Other HHS offices with responsibilities affecting the Head Start program include the Division of Payment Management, which disburses grant funds to grantees for HHS and other federal agencies, and the Division of Grants Policy, which develops policy and guidance for all discretionary grant programs administered by ACF.

[3]Early Childhood Programs: Local Perspectives on Barriers to Providing Head Start Services (GAO/HEHS-95-8, Dec. 21, 1994).

their families than do other early childhood centers; however, the quality of Head Start services has been uneven.[4] According to our third report, research that had been conducted was insufficient for determining whether Head Start has had a positive impact on participants' lives.[5] That report recommended that HHS include in its research plans an assessment of the impact of regular Head Start programs. HHS responded to this recommendation by stating that it would evaluate the feasibility of conducting impact studies. Our most recent report described participant characteristics, services, and funding.[6] We concluded, among other things, that the 33-year-old program is at a crossroads because the context in which it operates today differs greatly from that of 33 years ago when it was established. For example, Head Start's traditional part-day, part-year programs may not meet the needs of today's working families, and more providers are offering services today than when Head Start was established.

The Congress has recently acted to improve Head Start's and other federal programs' accountability. In passing the Government Performance and Results Act of 1993 (Results Act), the Congress sought to shift the focus of federal management from staffing, activity levels, and tasks completed toward results. The Results Act requires federal agencies to develop (1) strategic plans, (2) annual performance plans that establish goals and measures, and (3) performance reports for the Congress detailing progress made in meeting annual performance goals. In addition, in reauthorizing the Head Start program, the Head Start Act Amendments of 1994 required HHS to develop specific performance measures for the program to assess the achievement of desired outcomes. The Amendments also added a regulatory enforcement mechanism to the program. Grantees that do not meet time requirements for complying with program regulations now face termination of their funding. About 90 grantees have either relinquished their grants or been terminated since 1994. Although the majority of these grantees voluntarily relinquished their funding, ACF officials told us that they can often convince grantees to relinquish funding rather than face termination. To help HHS carry out its oversight responsibilities, the 1994 Amendments also required the agency, beginning in 1995, to provide the Congress with annual reports on grantees' compliance with program

[4]Early Childhood Centers: Services to Prepare Children for School Often Limited (GAO/HEHS-95-21, Mar. 21, 1995).

[5]Head Start: Research Provides Little Information on Impact of Current Program (GAO/HEHS-97-59, Apr. 15, 1997).

[6]Head Start Programs: Participant Characteristics, Services, and Funding (GAO/HEHS-98-65, Mar. 31, 1998).

regulations, including information on grantee efforts to correct program deficiencies. ACF plans to submit the first of these reports to the Congress in June 1998. ACF officials told us that workload demands prevented them from submitting these reports on time.

Head Start's Stated Mission, Goal, and Objectives Provide a Framework for Accountability

The Head Start Bureau has developed a performance assessment framework that reinforces program accountability by clearly linking the daily activities of local Head Start grantees to the program's overall strategic mission and goal. This framework not only emphasizes the importance of statutory and regulatory compliance, but also the importance of achieving demonstrable outcomes. Head Start developed this framework in response to congressional mandates and as a result of its own efforts. In particular, the Results Act and the Head Start Act Amendments of 1994 require the program to develop methods to measure its progress in meeting overall program purposes. In addition, the program's Advisory Committee on Head Start Quality and Expansion[7] recommended that a performance measurement system be developed to ensure that local Head Start grantees provide quality services to children and families. The Committee proposed that such a system include (1) identifying measurable outcomes, (2) selecting performance indicators and collecting indicator data, and (3) analyzing collected data. Although Head Start's framework is relatively new and largely untested, it is a major step toward being able to systematically demonstrate that local grantees comply with laws and regulations and that Head Start programs are collectively working toward achieving national program goals. It also improves Head Start's ability to measure the overall impact of the program on children's lives.

Head Start's Mission, Goal, and Objectives Focus Program Activities and Services on Results

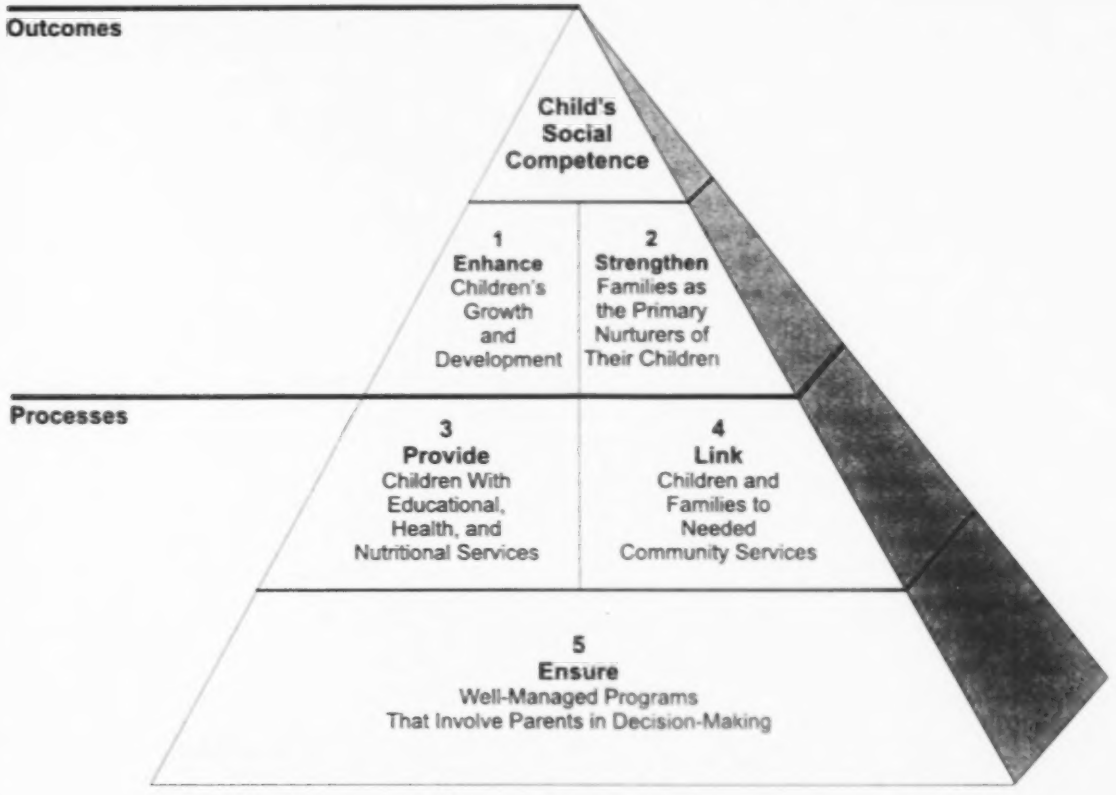
Head Start's strategic mission and goal are based on a philosophy that emphasizes the benefits of a comprehensive interdisciplinary program. Head Start enrollees receive educational, medical, nutritional, mental health, dental, and social services. Especially during Head Start's early years, the program provided services that participants probably would not have otherwise received. The program's mission and goal provide the context for the delivery of these services. Head Start's mission "to help prepare young economically disadvantaged children for success in school and life" provides the purpose for delivering these services; its goal to "improv[e] children's social competence" addresses the services' intended

[7]The Secretary of HHS created the Committee in June 1993 to review the Head Start program and make recommendations for improvement and expansion.

results. Head Start's strategic mission and goal reflect ACF's strategic goal of promoting the "healthy development, safety and well-being of children and youth." In turn, ACF's strategic goal reflects HHS' strategic goal of "improving the economic and social well-being of individuals, families and communities in the United States."

To guide program efforts toward meeting their purposes, the Head Start Bureau developed five performance-based objectives that reflect the different areas (such as cognitive, emotional, and social development; physical and mental health; and nutritional needs) that it believes make up social competence. The Bureau plans on using these five objectives to help assess, over time, its programs' quality, effectiveness, and results. As envisioned by the Results Act, Head Start's objectives shift the focus from the conduct of program activities toward a broader focus aimed at achieving results. Specifically, two of Head Start's five objectives focus on outcomes and are included in ACF's annual performance plan as annual performance goals. The remaining three objectives focus on program activities that the agency believes are critical to achieving its two outcome objectives. Figure 1 shows Head Start's five program objectives. It reflects the Bureau's belief that its five objectives form the foundation for improving children's social competence. The Bureau's two outcome objectives appear at the top of the pyramid. The Bureau's three objectives that focus on program processes or activities appear at the base of the pyramid.

Figure 1: Head Start Program's Strategic Goal and Objectives


Outcomes
Child's Social Competence
1 Enhance Children's Growth and Development
2 Strengthen Families as the Primary Nurturers of Their Children
Processes
3 Provide Children With Educational, Health, and Nutritional Services
4 Link Children and Families to Needed Community Services
5 Ensure Well-Managed Programs That Involve Parents in Decision-Making

Source: Head Start Bureau, ACF, HHS.

Head Start's Performance Measures Designed to Track Progress in Achieving Program Objectives

In developing the five program objectives, the Bureau sought to unify and organize performance measures that it developed to assess the quality and effectiveness of local grantees' Head Start programs. The 1994 Head Start Act Amendments require HHS to develop measures to assess local grantee program services and administrative and financial management practices annually and over longer periods. The measures are intended to help identify program strengths and weaknesses at both the regional and national levels. In addition, the Amendments directed the performance measures to be designed so that grantees could use them for performing

self-assessments and so that others could use them for conducting peer reviews of local programs.

Head Start developed a total of 24 measures to comply with the 1994 Amendments. The measures are designed to gauge the quality of health, educational, and nutritional services provided to enrolled children and assess local grantees' ability to manage their programs. Each of the 24 measures reflects one of the five program objectives. Like the objectives, the performance measures focus on either outcomes or processes and activities—outcome measures reflect outcome objectives, while process measures reflect process and activity objectives. For example, the Bureau developed three performance measures to assess progress toward its outcome objective of "strengthen[ing] families as the primary nurturers of their children." One of these performance measures is determining whether "Head Start parents demonstrate improved parenting skills." One of the Bureau's four performance measures for demonstrating progress in meeting the objective of linking children and families to needed community services is determining whether enrolled children's parents have secured child care so that these parents can pursue employment, education, or job training. In this case, both the objective and performance measure focus on local grantees' program activities. Table 1 shows the Bureau's 24 performance measures and their related program objectives.

Table 1: Head Start Program Performance Measures by Program Objective

Program objective	Performance measures
Enhance children's growth and development	Head Start children demonstrate improved 1. emergent literacy, numeracy, and language skills 2. general cognitive skills 3. gross and fine motor skills 4. positive attitudes toward learning 5. social behavior and emotional well-being 6. physical health
Strengthen families as the primary nurturers of their children	Head Start parents 7. demonstrate improved parenting skills 8. improve their self-concept and emotional well-being 9. make progress toward their educational, literacy, and employment goals
Provide children with educational, health, and nutritional services	Head Start 10. programs provide developmentally appropriate educational environments 11. staff interact with children in a skilled and sensitive manner 12. programs support and respect children's cultures 13. children receive needed medical, dental, and mental health services 14. children receive meals and snacks that meet their daily nutritional needs 15. programs provide individualized services for children with disabilities
Link children and families to needed community services	Head Start parents 16. link with social service agencies to obtain needed services 17. link with educational agencies to obtain needed services 18. link with health care services to obtain needed care 19. secure child care in order to work, go to school, or gain job training
Ensure well-managed programs that involve parents in decision-making	Head Start 20. programs are well managed 21. parents are involved actively in decisions about program operations 22. programs employ qualified staff 23. programs support staff development and training 24. programs comply with Head Start regulations

For each of the 24 measures, the Bureau has identified indicators and data sources. As used in Head Start's performance assessment framework, indicators are statements that relate how data will be used in relation to a particular performance measure. The Bureau has developed from one to seven indicators for each measure. For example, the Bureau uses three indicators to determine whether participating children's parents have secured child care so that these parents can pursue employment, education, or job training. The indicators are the (1) number and

percentage of programs that provide child care, (2) number and percentage of programs that provide linkages to child care, and (3) number and percentage of program children's parents that report having stable child care services. To obtain data for each indicator, the Bureau plans on relying on its management information systems or initiating new data collection efforts. In some cases, the Bureau will collect data from each Head Start grantee; in other cases, the Bureau will use only a sample of grantees to obtain needed information.

Local Head Start Agencies Must Comply With Program Regulations

Performance standards (hereafter referred to as program regulations) are the regulations that define local program activities; grantees must adhere to these regulations in operating their programs. According to the Bureau, the regulations consistently define the quality of services that grantees must deliver and constitute the single most important statement of the expectations and requirements that grantees must meet. HHS' program regulations generally reflect Head Start's performance measures, according to our review.

HHS ensures local program quality by monitoring and enforcing compliance with these regulations, Head Start officials said. According to HHS, grantee compliance with the agency's program regulations forms the foundation for obtaining positive program outcomes. The Bureau plans on demonstrating its progress in meeting its objectives by analyzing, among other things, compliance information it gathers during on-site inspections of local Head Start grantees and by conducting national surveys of Head Start children and their families.

Implementation of System for Ensuring Regulatory Compliance Could Be Improved

Head Start uses several processes to assess and enforce local Head Start agencies' compliance with program regulations. On-site inspections are the main enforcement mechanism. The law requires HHS to conduct such an inspection after a grantee's initial operating year and at least once during each 3-year period thereafter. Conducting on-site inspections not only allows HHS to observe local program operations first hand, but also allows it to collect other data by examining grantee records such as children's immunization schedules. HHS has tailored an approach to conducting inspections that is intended to result in a comprehensive review of local program operations; concerns have been raised, however, about the consistency of its implementation. Inspections may not be conducted consistently among regions and by inspectors due to the subjectivity involved in conducting inspections and the inspectors' levels

of expertise, according to our review. HHS has tried to improve the consistency of its inspections, but the impact of these efforts is still unclear.

On-Site Inspections Are Designed to Comprehensively Assess Grantee Compliance With Program Regulations

To conduct on-site inspections, HHS uses Head Start program specialists from ACF regional offices located nationwide. These staff, in addition to peer reviewers, including local Head Start agency directors and consultants with particular expertise, inspect Head Start programs operating in their respective regional office's jurisdiction. The use of expert reviewers who have different specialties, such as education or nutrition, ensures that knowledgeable individuals are collecting data on the different Head Start program components. On-site inspections generally take 1 week to complete and conclude with a meeting among inspection team members and grantee managers and staff to discuss the team's observations and findings.

Team members use a specially designed on-site program review instrument (OSPRI) to conduct inspections. The instrument helps ensure that team members conduct a comprehensive assessment of grantees' operations. It reflects Head Start's program objectives and covers such program components as education, health and social services, program administration and finances, parental involvement, and facilities. Within each of these components, the instrument lists review items, or criteria, that team members use to assess grantees' operations. Team members must assess grantee operations using a total of 256 criteria covering all program areas.[8] Team members record their individual judgment on the OSPRI about whether a grantee has met each criterion; they must judge grantees as either "in compliance" or "not in compliance" for each criterion. Table 2 lists program components on the OSPRI and selected related criteria for assessing compliance.

[8]ACF is revising the version of the OSPRI that was in use during our review. The revised version will be based on the new performance standards, which became effective in Jan. 1998. ACF expects to complete its OSPRI revision by fiscal year 1999; until then, ACF is using an interim version. Criteria used to assess fiscal compliance will not change from the original version of the OSPRI, according to ACF officials.

Table 2: Program Components on the OSPRI and Selected Related Criteria

Component	Criteria
Education	There is a supportive social and emotional climate.
Health	A health services advisory committee helps plan, operate, and evaluate the health services program.
Mental health	A mental health professional is available to the program and to children.
Nutrition	There is a written nutrition plan, annually updated.
Social services	There are established procedures for emergency assistance and crisis intervention.
Parent involvement	Parent training/orientation is provided to prevent child abuse and neglect and to protect abused and neglected children.
Disabilities services	Resources to implement the disabilities services plan are adequate.
Eligibility and recruitment	The program maintains a waiting list that ranks children by the program's selection criteria to assure children enter the program as vacancies occur.
Administration	The grantee has and implements a written procedure that assures specific program objectives and activities are completed in a timely manner.
Staffing requirements and program options	The grantee adequately supervises its staff.
Financial and property management	The grantee has written accounting procedures.

On the basis of team members' individual judgments, the Head Start program specialist must decide whether a grantee is in compliance with program regulations. If the specialist decides that a grantee is not in compliance, then a judgment on the severity of noncompliance must also be made. Under program guidelines, a grantee that is not in compliance with regulations must be classified as either (1) noncompliant or (2) deficient. Noncompliant grantees are those that have not complied with regulations in one or more program components. For example, a grantee may be classified noncompliant with the health component because children's dental or medical records are incomplete. Deficient grantees are those that have not complied with regulations and whose failure to comply impedes the grantee's ability to provide quality services to children and families. For example, a grantee may be classified deficient if its program has a staffing problem that affects the education, social, and health services provided to children.

A noncompliant grantee must correct identified problems within 90 days and notify its ACF regional office that it has done so. A grantee classified as

deficient, however, faces far more severe consequences than a grantee classified as noncompliant. A deficient grantee must prepare a quality improvement plan describing how it will correct all identified deficiencies and submit it to the respective ACF regional office for approval. The grantee has up to 1 year to correct identified deficiencies, during which HHS makes training and technical assistance available to the grantee to help correct deficiencies. If the grantee does not correct all deficiencies within 1 year, the ACF regional office must terminate the program. Regional ACF staff monitor grantees to ensure that they take corrective action.

ACF staff use a national database called the Head Start Management Tracking System to help them monitor grantee compliance. Head Start regional office staff enter information gathered from their on-site inspections into the database, including the compliance status of grantees regarding the 256 criteria used to conduct the inspections.

On-Site Inspections Could Be More Consistent

Both ACF regional staff and outside researchers have raised concerns about the consistency of on-site inspections. In 1993, a study prepared under contract for ACF noted wide variation among regions in the number of the OSPRI items for which grantees were judged as out of compliance.[9] The researchers concluded that this wide variation was probably not caused solely by differences in grantees' activities, but also stemmed at least partially from differences in the way regional office staff judged compliance. A study in 1996 by the same contractor also identified ensuring consistency in interpreting inspection results as a major challenge for Head Start.[10] In our interviews with ACF regional staff, they also expressed concern that on-site inspections may not be conducted consistently among regions and by reviewers.

Several factors probably contribute to this inconsistency. First, some of the variation may derive from the OSPRI. For example, the OSPRI includes somewhat subjective items, such as "the parent involvement program adequately provides methods and opportunities for involving parents in experiences in child growth and development" and "the grantee provides adequate supervision of its staff" as well as more objective questions such as "completion of all recommended immunizations." Second, the lack of guidance on the relative importance of specific OSPRI items may further obscure the inspection process. Third, some of the variation may result

[9]Indices of Head Start Program Quality, Pelavin Associates, Inc. (Washington, D.C.: Sept. 1993).

[10]Revision of the Head Start Federal Monitoring System, Pelavin Research Institute (Washington, D.C.: Jan. 3, 1997).

from differences in inspectors' expertise. In the 1996 study, some focus group participants stated that inspectors without the necessary expertise are sometimes selected. Other factors, such as the time and work necessary to monitor (and, if necessary, terminate) poor performers, may discourage inspectors from classifying grantees as deficient.

According to ACF officials, the inspection process includes controls to help ensure consistent inspections and ACF has taken several steps to improve the process. For example, to help lessen the subjectivity of judgments about compliance, part of the inspection process includes team meetings during which inspectors can discuss their opinions about grantee performance. In addition, training is provided to inspectors before they conduct an on-site inspection, and ACF sometimes uses consultants with particular expertise. ACF also hosts annual conferences where regional ACF staff can share best practices, identify training needs, and discuss other ways to increase the consistency of inspection teams' judgments. ACF is also trying to clarify priorities regarding grantee compliance in its revision of the OSPRI. Instead of concentrating on assessing agencies' compliance with individual program regulations, ACF is adopting an approach in which inspectors would focus on reviewing the systems a grantee uses to implement regulations. Inspectors would then assess compliance with individual regulations to the extent that inspectors find problems using this systems approach. The impact of these efforts, however, is still unclear.

Although the full impact of the inconsistency of on-site inspections is unknown, inconsistent inspections could lead to uneven treatment of grantees. In addition, ACF officials told us that they use data from the Head Start Management Tracking System—which includes information from on-site inspections—to help make program decisions and prepare congressionally mandated reports. Use of data from the tracking system, therefore, may not be as valuable as it could be to the Congress and ACF in making decisions about Head Start program policy.

Other Processes Used to Monitor Grantee Regulatory Compliance Also Have Limitations

In addition to conducting on-site inspections, regional ACF staff also monitor grantees' compliance with regulations by annually reviewing their financial audit reports and their grant applications.[11, 12] This monitoring supplements the triennial on-site inspections. In general, financial audits are not designed to provide detailed assessments of grantees' compliance with all program regulations. Rather, financial auditors generally determine whether grantees have established and implemented internal controls that function to help ensure that programs comply with regulations. Auditors may select and review samples of financial transactions to determine whether a grantee has followed established procedures and program regulations. If a grantee administers more than one federal grant, as is often the case with large nonprofit agencies, school districts, and municipalities, relatively small grants may not be reviewed in as much detail, if at all, as larger grants. As a result, Head Start programs that constitute only a small portion of a grantee's total federal funding may not be reviewed or may not have transactions related to the program selected for compliance review.

In addition to a lack of detail, financial audit reports may not provide timely information for monitoring current grantee operations. Grantees have 9 months to submit financial audit reports for any given year.[13] It may take several additional months before officials in the HHS Office of Inspector General review the audit report, summarize findings, and submit the information to regional ACF staff, who monitor grantees' resolutions of audit findings. Grantees that fail to resolve audit findings within 6 months may be classified as "high risk" and be required to submit additional financial reports or be subject to additional financial controls.[14] However, high-risk grantees generally do not face termination of their funding unless they are also classified as deficient, which usually involves an on-site inspection. As a result, ACF may wait up to 3 years until the next regularly

[11]In general, the Single Audit Act requires that grantees that receive federal funding undergo periodic financial audits generally on an annual, or in certain cases, biennial basis. An independent auditor conducts these audits, which include a review of grantees' accounting and financial management systems.

[12]ACF regional staff also review other financial reports submitted by grantees, including federal cash transaction reports and financial status reports. The cash transaction reports provide data on grantees' federal cash balances, while the status reports provide aggregate data on grantees' expenditures. Late submissions of these reports indicate that a grantee may be having difficulty complying with program regulations and may highlight a need to more closely monitor the grantee, according to several regional staff.

[13]Before June 30, 1997, a federal grantee had 13 months from the end of its grant year to submit an independent financial audit for that year.

[14]For example, a grantee may only be allowed to receive funds on a cost-reimbursable basis or be required to provide documentation to support program expenditures.

scheduled triennial inspection before it classifies a high-risk grantee deficient and requires it to develop a quality improvement plan and face termination from the program. ACF officials encourage regional staff to conduct inspections, they said, before regularly scheduled triennial inspections to determine whether high-risk grantees should also be classified deficient.

Another tool for monitoring compliance is the review and approval of Head Start grant applications and program budgets, which is done by ACF regional staff. Grantees are generally required to submit a noncompetitive grant application annually, including a proposed budget. Regional staff told us that late applications for refunding may indicate that a grantee is having difficulty complying with program regulations. Regional staff cannot compare budgets with actual expenditures at the needed detailed level, however, because grantees do not have to report detailed expenditure data.

Head Start Processes Improving but Still Insufficient to Assess Whether Program Purposes Achieved

HHS has taken significant steps toward developing a comprehensive strategy to assess whether the Head Start program is achieving its purposes. Initially, HHS plans to measure program outcomes using the results from a nationwide survey of a representative sample of children enrolled in Head Start programs and their families. HHS will use survey results to project the degree to which it is accomplishing program purposes on a national level. In the future, HHS plans to assess whether Head Start is achieving its purposes by measuring individual grantees' progress in achieving desired outcomes, rather than relying on a sample of grantees. HHS believes this strategy is also sufficient to determine program impact. Although the strategy will move HHS closer to being able to make such a determination, HHS could do more to ensure that it accurately measures the program's actual impact with a greater degree of confidence. HHS would incur costs in such an effort; however, the significant financial commitment the federal government has made to the Head Start program warrants such an investment.

Initiatives Will Provide Framework for Assessing National Outcomes but Not Local Agency Outcomes

Overall, HHS has a methodologically and conceptually sound approach to assessing outcomes. It has designed initiatives to assess its progress in meeting its two outcome-focused objectives: (1) enhancing children's growth and development and (2) strengthening families as the primary nurturers of their children. HHS has developed multiple performance measures to use in assessing progress in meeting these objectives. For

each measure, HHS has established one or more performance indicators that will track the percentage of change from the previous year.

Because data on these indicators were not available, HHS has developed a strategy to obtain them, namely, the Family and Child Experiences Survey (FACES), a study of a representative sample of Head Start children and their families designed to generate national-level data. FACES will provide most of the outcome data needed to assess whether outcome-focused objectives are being met on an ongoing basis. According to HHS officials, a spring 1997 FACES pilot resulted in collecting data from a sample of 2,400 families with children enrolled in 160 randomly selected centers in 40 Head Start programs nationwide. Subsequently, the full study has involved data collected in the fall of 1997 and spring of 1998. The overall design is to collect data on 3,200 children and their families at program entry, exit (or completion of each year of Head Start), and at the end of kindergarten. In conducting the study, researchers plan on using well-established and widely used scales, assessments, and observational checklists and specially tailored questionnaires. These instruments will collect data on children's vocabulary, literacy and mathematical skills, perceptual-motor development, and social and communicative competence as well as information about the families.

By using well-established instruments, HHS will be able to assess specific outcomes, such as literacy improvement or expressive language. Meanwhile, HHS' use of a variety of instruments and scales will facilitate assessing many dimensions of cognitive, emotional, and physical growth and development. Its sampling approach is designed to provide a national perspective on how well Head Start children and families are doing. Furthermore, HHS has ensured that the performance measures in its strategic framework are represented in the national survey.

Although HHS has focused on outcomes at the national level, it has not focused on outcomes at the local agency level, even though many Head Start agencies receive substantial funding. In fiscal year 1997, funding ranged from about $120 million in New York City to about $135,000 for one grantee in Minnesota, with 66 percent of the local agencies receiving grants of at least $1 million. Grantees and their respective delegate agencies, however, do not have to use the measures associated with the two outcome-focused objectives to assess whether outcomes have been achieved. Instead, according to ACF officials, the performance measures are intended at this time to provide a self-assessment tool for individual programs for strategic planning purposes. Later on, ACF intends to require

grantees to demonstrate their progress in achieving outcomes, according to ACF officials. They indicated that the transition from grantees' complying with standards to assessing outcomes is a challenging one, and they expect it will take time to work with the grantees to enable them to assess outcomes. Assessing local programs' outcomes will be difficult, according to HHS, because each agency operates in a unique community. According to ACF, some grantees are already assessing outcomes and ACF plans to use findings from their experiences to help other grantees. ACF has not yet developed a plan or a timetable, however, for this transition.

HHS Initiatives Will Not Adequately Assess Program Impact

According to HHS, its strategy for assessing outcomes will also allow it to determine Head Start's impact on children's growth and development. HHS' plans will not provide the needed information about program impact, however, because they do not include a research study or set of studies that will definitively compare the outcomes achieved by Head Start children and their families with those achieved by similar non-Head Start children and families.

HHS Plans Many Comparisons to Draw Conclusions About Impact

HHS has identified comparisons it believes will provide a basis to draw conclusions about Head Start's impact on children's developmental competencies. HHS will rely mainly on the FACES initiative for making these comparisons. HHS has acknowledged, however, that the absence of a control or comparison group makes it difficult to explicitly attribute any participant progress to the Head Start program. HHS proposes, as a solution to this problem, to compare FACES data with existing data on the national population of children of the same age and those from low-income families who have not attended Head Start.

Some of the assessment tools used in FACES are nationally normed instruments, which is why HHS is relying on them for its comparisons with children in the national population.[15] By using these instruments, HHS can compare the FACES results with the average scores of all children of the same age. As Head Start collects data on children at the end of each year of Head Start participation and at the end of kindergarten, it will be able to compare the rate of development of Head Start children with all preschoolers of the same ages. For example, even though Head Start children may be behind their age mates developmentally, they may show equivalent amounts of growth between the fall and spring of the Head Start year or from 1 year to the next.

[15]Norms are obtained by administering a test to a sample of people and deriving the distribution of scores for that group. Some of the tests used by Head Start have been normed using samples selected to represent the national population for a particular age group.

Another comparison HHS has identified would infer Head Start's impact from differences in children who participated in the program to different degrees or for varying durations. For example, if children who attended Head Start classes for many hours a week instead of a relatively few hours or for 2 years instead of 1 year showed greater developmental gains, then the difference could be evidence of Head Start's impact.

In addition, HHS has described comparing children who have attended Head Start with those who have not as part of its effort to assess impact. Data on the Head Start children would come from FACES; data on non-Head Start children would come from the National Center for Education Statistics' (NCES) National Household Education Survey (NHES) on the developmental status and preschool program attendance histories of a sample of children. HHS could not compare these children using any measure that requires direct observation by a trained interviewer or teacher, but it would be able to compare parent reports of children's developmental accomplishments and difficulties.

Another comparison option for HHS, which would not necessarily include data from the FACES study, involves NCES' Early Childhood Longitudinal Study (ECLS). HHS is collaborating with NCES on this study, implemented in fall 1997 after a 5-year planning effort. The study will collect data on a nationally representative sample of kindergarten children in public and private schools. The study will collect data from parents and children, including descriptions of children's preschool experiences and performance on standardized tests in areas such as achievement and psychomotor development. Although information on the early childhood educational experience of children will be limited, according to HHS, it will be possible to compare the school progress of children who had attended Head Start before entering the study with that of low-income children without a center-based early childhood education experience.

Likely Differences Among Children in Comparisons Limit HHS' Interpretations

The noncomparability of children in the comparison groups will limit HHS' ability to assess Head Start's impact. Without comparable groups, outcomes could be mistakenly attributed to Head Start participation when these outcomes were really caused by other factors related to differences in the comparison groups.

HHS' plans to compare FACES results with national norms are limited to the extent that the norms are not based on children like those in Head Start programs but are generally based on children more representative of the national population. Thus, if the children in the group used to establish the

norms are unlike the children in Head Start, conclusions about program impact will be unclear. Furthermore, comparing norms over time to assess gains assumes that the groups have an equivalent learning rate, which may not be the case. The lack of assurance that groups are comparable similarly limits HHS' plans to compare selected data with the NHES and ECLS samples. Although comparing Head Start results with other study results moves in the direction of assessing impact, it is not definitive enough.

Given the size of the Head Start program and plans for its expansion, investing in studies that will definitively assess its impact is important. As HHS has acknowledged, the most conclusive way to determine program impact is to compare a group of Head Start participants with an equivalent group of nonparticipants. The preferred method for establishing the equivalency of groups at the outset is to randomly assign participants to either a Head Start group or a comparison group. No matter how extensive the efforts to document the equivalency of groups formed in other ways, the groups' equivalency will remain uncertain. If the groups are not equivalent, outcomes attributed to Head Start will be in doubt. For instance, a recent evaluation of the Comprehensive Child Development Program,[16] a demonstration project involving comprehensive early childhood services like Head Start's, found positive changes in participating families. The study compared participants with comparable nonparticipants, and researchers discovered that nonparticipating families had positive changes similar to the participating families. They concluded, therefore, that the positive changes could not be attributed to program participation. Researchers in this study had confidence that the groups were comparable when the participants entered the program because participants had been randomly assigned to groups.

To obtain information about impact, HHS could conduct a study or set of studies using random assignment of Head Start-eligible children and their families. HHS could design studies to compare outcomes of Head Start participants with those of nonparticipants. Nonparticipants would not necessarily receive no services but could participate in other educational settings, child care, or in any other programs. HHS could also explore the impact of different Head Start approaches used by different programs by comparing the outcomes in these local Head Start programs with each other as well as with other non-Head Start programs.

[16] National Impact Evaluation of the Comprehensive Child Development Program. Abt Associates (Cambridge, Mass.: June 1997).

Conclusions

Head Start has, through the years, provided a comprehensive array of services and, as prompted by the Results Act, has in recent years substantially strengthened its emphasis on determining the results of those services. Furthermore, recent legislation has established significant consequences for Head Start grantees that do not comply with requirements and placed new requirements on agencies to assess achievement of program purposes. The Head Start program's performance in meeting challenges for these new requirements is uneven, however.

HHS has established a comprehensive process to ensure that Head Start grantees comply with program regulations and has recently taken steps to address weaknesses in its on-site inspection process to ensure the consistency of inspections. Given that on-site inspections are HHS' main mechanism for ensuring regulatory compliance, the integrity of this process is essential to an effective enforcement program.

In addition, improving the consistency of this process would give HHS more confidence in the information management systems that track these data. In turn, HHS would be in a better position to provide the Congress with the information and reports it requested when the Congress last reauthorized Head Start. These data can also provide valuable information at the national level that would be useful for developing policy and practices for guiding grantees toward improved compliance and service delivery.

Head Start's initiatives, in particular, FACES, are headed in the right direction because of their increased focus on outcomes and research that could be expanded to compare outcomes for children in Head Start with those for similar nonparticipating children and families. HHS' current processes, however, focus only on national-level outcomes based on data collected through nationally representative samples. Although this type of outcome assessment would be new and therefore challenging to grantees, knowing whether local programs achieve results is important. Given the significant federal investment in the Head Start program, collecting such data is worthwhile.

Head Start's planned processes to assess program impact—that is, whether a difference in outcome has resulted from Head Start participation—do not provide a definitive assessment of the program's overall impact. The federal government's significant financial investment in the Head Start program, including plans to increase the number of children served and enhance the quality of the program, warrants definitive research studies, even though they may be costly. HHS has no

plans for a research study or set of studies that will definitively compare the outcomes achieved by Head Start children and their families with those achieved by similar non-Head Start children and families. Consequently, questions about Head Start's impact will remain unanswered.

Recommendations to the Secretary of the Department of Health and Human Services

To ensure that individual Head Start grantees are held accountable for achieving program purposes, we recommend that HHS develop and implement a plan for assessing individual grantees' performance based on their achieving the outcomes associated with HHS' performance objectives. Such a plan could include, for example, guidance and suggested methods for grantees to use in assessing the degree to which children show improvement in critical outcome areas such as cognitive skills, literacy, and gross motor skills.

To determine whether the Head Start program is making a difference in the lives of those it serves, we recommend that HHS assess the impact of regular Head Start programs by conducting a study or studies that will definitively compare the outcomes achieved by Head Start children and their families with those achieved by similar non-Head Start children and families.

Agency Comments and Our Evaluation

In commenting on a draft of our report, HHS affirmed its commitment to accountability and results for the Head Start program. It agreed with our recommendation that a plan be developed and implemented to assess individual grantees' performance but noted that such a plan should be developed cautiously to ensure that the measures have no unintended consequences such as diverting programs from their historic commitment to serving a community's most disadvantaged children. We agree that such plans should be developed carefully and in consultation with experts. Meanwhile, a timetable should be established that delineates when and how grantees would be held accountable and when HHS will provide guidance to grantees on types of measures and instruments.

Regarding our recommendation that the Secretary conduct a study or set of studies to assess the impact of regular Head Start programs, the Department expressed the belief that its current research strategy of multiple studies and comparisons will produce sufficient information about impact. We agree with HHS that its strategy of multiple studies and comparisons will provide useful information about outcomes achieved by

Head Start children and that the comparisons with non-Head Start children can be used to draw some limited conclusions about Head Start's impact. HHS has acknowledged that these conclusions will have limitations, however, because of the study designs. We believe HHS' disagreement with our recommendation reflects in part a misunderstanding of its nature and how it could be implemented. HHS' comments imply that we recommended that its only research activity should be a single, nationally representative impact study that would assess the impact of Head Start by comparing children in Head Start with children receiving no early childhood education services. We are not now nor have we ever recommended that the Department abandon all other research activities to focus exclusively on a single study to assess impact. Instead, we recommended in an earlier report that the Department include in HHS' research plan an assessment of the impact of regular Head Start programs, and we are now recommending that HHS conduct a study or studies, along with its other research efforts, that will definitively assess the program's impact.

HHS also expressed doubt that studies using random assignment of children to Head Start to ensure that Head Start participants are comparable with nonparticipants could feasibly be done and that such studies would provide more definitive information about the program's impact. HHS has acknowledged, however, that it is working with local partners to determine the feasibility of implementing studies that use a random assignment design to assess Head Start's effectiveness.

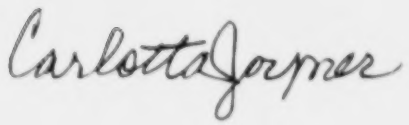
We disagree with HHS' conclusion that designs using random assignment are not feasible. Although we acknowledge the difficulty of designing and implementing studies with random assignment, some of the ethical, methodological, and logistical difficulties that HHS cites reflect in part several questionable assumptions about how we would expect such studies to be done. First, HHS notes that mandates to serve the most needy of eligible families and to incorporate a minimum percentage of children with disabilities guide Head Start's recruitment. HHS cites the idea of recruiting more children than would be served using a random assignment study design as a difficulty in implementing the recommendation. It is not clear how or why this would be an unsurmountable problem. A second difficulty cited—apparently a logistical one—is finding enough eligible children not in Head Start. Relevant to both of these issues is a point HHS has frequently made in other documents: that large numbers of children are eligible for Head Start but are not being served. Plans to expand the program are based on this assumption of unmet need. To the extent that a significant unmet need for Head Start exists, then using a lottery to

determine which among the many needy children would receive the service should not be unsurmountable. A third issue hinges on HHS' assertion that the kind of study we recommended would require that the children not participating in Head Start would have to receive no early childhood education services at all or exactly identical non-Head Start services. A study using randomization could have more than one comparison group. In fact, by using multiple alternative comparison groups, HHS could compare outcomes achieved by Head Start participants with outcomes of children in other programs. Furthermore, children do not necessarily have to receive no services to be part of a comparison group; that is, families in a comparison group could choose to participate in child care settings or other preschool programs.

Finally, HHS maintained that the mobility of low-income populations would probably result in children leaving treatment in unpredictable numbers and at unpredictable points in time or crossing among program and comparison groups. Although attrition and mobility can occur, a well-designed study would incorporate these kinds of factors and develop mechanisms to account for their effect. For example, follow-up studies would allow researchers to obtain information that would allow them to account for the extent of attrition or mobility. Furthermore, collecting information on attrition would be an interesting characteristic to examine for determining the similarity of attrition patterns.

Despite our areas of disagreement with HHS, we are encouraged that the agency is continuing to study the feasibility of conducting studies that include random assignment and is willing to consider seeking additional advice from experts in research design. We hope that the results of these activities will lead to studies producing more definitive conclusions about Head Start's impact than can be drawn from the studies currently planned. Such studies would be extremely useful additions to Head Start's considerable body of ongoing research for assessing and improving the program. The full text of HHS' comments appears in appendix I.

We are sending copies of this report to the Secretary of Health and Human Services, appropriate congressional committees, and other interested parties. Please call me at (202) 512-7014 or Harriet Ganson, Assistant Director, at (202) 512-9045 if you or your staff have any questions about this report. Major contributors to this report are listed in appendix II.



Carlotta C. Joyner
Director, Education and
Employment Issues

List of Requesters

The Honorable Dan Coats
Chairman, Subcommittee on Children and Families
Committee on Labor and Human Resources
United State Senate

The Honorable Christopher J. Dodd
Ranking Minority Member
Subcommittee on Children and Families
Committee on Labor and Human Resources
United States Senate

The Honorable Bill Goodling
Chairman, Committee on Education and the Workforce
House of Representatives

The Honorable William Clay
Ranking Minority Member
Committee on Education and the Workforce
House of Representatives

The Honorable Frank Riggs
Chairman, Subcommittee on Early Childhood, Youth and Families
Committee on Education and the Workforce
House of Representatives

The Honorable Matthew G. Martinez
Ranking Minority Member
Subcommittee on Early Childhood, Youth and Families
Committee on Education and the Workforce
House of Representatives

The Honorable Dale Kildee
House of Representatives

The Honorable Randy Cunningham
House of Representatives

Contents

Letter 1

Appendix I
Comments From the Department of Health and Human Services 30

Appendix II
GAO Contacts and Staff Acknowledgments 42

Tables

Table 1: Head Start Program Performance Measures by Program Objective 10
Table 2: Program Components on the OSPRI and Selected Related Criteria 13

Figure

Figure 1: Head Start Program's Strategic Goal and Objectives 8

Abbreviations

ACF	Administration for Children and Families
ECLS	Early Childhood Longitudinal Study
FACES	Family and Child Experiences Survey
HHS	Department of Health and Human Services
NCES	National Center for Education Statistics
NHES	National Household Education Survey
OSPRI	on-site program review instrument

Comments From the Department of Health and Human Services

DEPARTMENT OF HEALTH & HUMAN SERVICES

Office of Inspector General

Washington, D.C. 20201

JUN 22 1998

Ms. Carlotta C. Joyner
Director, Education and
Employment Issues
United States General
Accounting Office
Washington, D.C. 20548

Dear Ms. Joyner:

Enclosed are the Department's comments on your draft report, "Head Start: Challenges in Ensuring Effective Program Implementation and Demonstrating Results." The comments represent the tentative position of the Department and are subject to reevaluation when the final version of this report is received.

The Department appreciates the opportunity to comment on this draft report before its publication.

Sincerely,

Michael Mangano
for June Gibbs Brown
Inspector General

Enclosure

The Office of Inspector General (OIG) is transmitting the Department's response to this draft report in our capacity as the Department's designated focal point and coordinator for General Accounting Office reports. The OIG has not conducted an independent assessment of these comments and therefore expresses no opinion on them.

Comments of the Department of Health and Human Services on the U.S. General Accounting Office's Draft Report, "Head Start: Challenges in Ensuring Effective Program Implementation and Demonstrating Results," GAO/HEHS-98-186

General Comments

We appreciate the opportunity to comment on the General Accounting Office (GAO) draft report and thank GAO for the collaborative approach that was used to develop the report.

We share GAO's commitment to accountability and results. We know that Head Start works and we take seriously our responsibility to ensure that all Head Start programs are held accountable for providing services of the highest quality. We continue to work hard toward that end. We appreciate GAO's recognition that "HHS has developed a performance assessment framework that reinforces Head Start program accountability by linking specific program activities with the program's overall strategic mission and goal." and that "...the Head Start program has put itself in the position of being able to answer questions about whether its program is complying with regulations and whether the program is achieving its purposes."

Because accountability and results are so central to our work, the Department has made tremendous progress in moving the Head Start program forward since the enactment of the 1994 reauthorization of the Head Start Act. Through a broad consultation effort, we have revised the core set of regulations that define quality Head Start services, the Performance Standards. As directed by the Act, we have developed a set of Performance Measures that measure, annually and over longer periods, the quality and effectiveness of programs operated by Head Start agencies. We have developed a conceptual framework to display linkages between process and outcome measures for Head Start children and families and have begun, for the first time, collecting child outcome data through the Head Start Family and Child Experiences Survey (FACES) from a nationally representative sample of Head Start programs, classrooms, teachers, parents and children. We have addressed concerns related to poor grantee performance by implementing the portion of the Act that directs us to take specific steps to correct grantee deficiencies once they are identified and to begin termination proceedings for grantees that fail to take corrective action. To date, 90 grantees have been terminated of have chosen to relinquish their Head Start grant because of poor performance. We have established a new program for infants and toddlers called Early Head Start and put in place a proactive research agenda. We have revised the training and technical assistance system. And we have undertaken a critical review of the monitoring process. While we believe that we have a strong and comprehensive approach to reviewing grantees, we are revising the system and engaging in a continuous improvement process.

We have provided a more detailed account of our accomplishments in the areas related to program monitoring and research and evaluation beginning on page 7 of this response.

While we have made great progress, we know that there are further steps to be taken. While we share many of GAO's views about what those next steps should be, we have some technical reservations about the report's recommendations based on the best advice of academics and experts with whom we have consulted extensively. In the area of research, the report's specific proposal for randomized control groups in a national study is in conflict with the recommendations of a wide range of academic researchers. In the area of monitoring, we share the goal of adding an outcome focus for individual programs to our existing strong process of monitoring quality and the particular elements of quality that we know lead to outcomes, but believe that creating a national strategy for incorporating this focus into monitoring needs to be done thoughtfully and carefully. Experts in the development of young children continue to emphasize to us how important it is to move step-by-step on outcome measures for your children in order not to have unintended consequences, such as moving programs away from their historic commitment to serving the most disadvantaged children in a community.

In our role as a national laboratory for services to young children, every step that we take with the Head Start program is intended to strengthen the quality of services to young children and their families and to develop new knowledge to inform the entire early childhood field on what are best practices for young children. Our responses to the recommendations in this report reflect this approach.

GAO Recommendation

To ensure that individual Head Start grantees are held accountable for achieving program purposes, we recommend that HHS develop and implement a plan for assessing individual grantees' performance based on their achieving the outcomes that are associated with HHS' performance objectives. Such a plan could include, for example, guidance and suggested methods for grantees to use in assessing the degree to which children demonstrate improvement in critical outcome areas such as cognitive skills, literacy and gross motor skills.

Department Comments

As GAO and others have noted, Head Start's monitoring strategy involves in-depth assessment of program quality, including those elements of quality that are most lined to results. We believe that the strength of our monitoring strategy is one of the reasons for the high quality that exists in so many Head Start

programs. We are in the process of revising our monitoring approach and GAO's suggestions regarding consistency and an outcome focus are consistent with the work that is underway.

We do want to note that there are a number of reasons for proceeding step-by-step and with some caution when moving toward outcome measures for individual programs for young children. Since one of the consequences of monitoring could be termination of Head Start funding, we need to be very careful about what we are holding grantees accountable for. We also need to be clear that we are measuring the most appropriate areas with appropriate instrumentation. Since we are breaking new ground with outcome measures for the early education of children, we need to learn more from the Performance Measures effort about appropriate outcome measures for young children. We are learning a great deal from the FACES methodology about what works but we need to know more. We need to engage in a discussion with a variety of experts in the field of early childhood education and tests and measurements from the academic, research and practitioner communities regarding the best approach for using outcome measures in monitoring all grantees.

GAO Recommendation

To determine whether the Head Start program is making a difference in the lives of those it serves, we recommend that HHS assess the impact of regular Head Start programs by conducting a study or studies that will definitively compare the outcomes achieved by Head Start children and their families with those achieved by similar non-Head Start children and families.

Department Comments

We agree that it is vital that the Department undertake research to examine the effects of Head Start on children and families. To this end, we have adopted a multi-dimensional research strategy which combines a comprehensive national study of Head Start programs, children and families with comparisons to a variety of national data sets; collaborative linkages with other large national studies; and includes smaller, locally-based feasibility studies for determining Head Start effectiveness.

There is a large body of convincing research on the short and intermediate effects of Head Start that demonstrates benefits to children and families. For example, a report by the Packard Foundation's Center for the Future of Children (1995), based on the results of nearly 150 studies of Head Start and other preschool programs, concludes that children exposed to early childhood education have higher reading scores, are less likely to be held back a grade or to be placed in expensive special

education classes.

The GAO recommends that the Department should conduct further research to determine whether the Head Start program is making a difference in the lives of those that it serves. The report praises the Department for its "...methodologically and conceptually sound approach to assessing outcomes." The GAO recommends that the Department should "...assess the impact of regular Head Start programs by conducting a study or studies that will definitively compare the outcomes achieved by Head Start children and their families with those achieved by similar non-Head Start children and families." We are convinced that the current course of Head Start quality research will result in such comparisons.

Since the 1994 reauthorization, Head Start has created a wealth of new initiatives and systems to promote stronger accountability and program quality. In 1995, the Head Start Bureau established a Quality Research Consortium (QRC). One objective of the QRC is to work cooperatively to develop and test instruments to gather information on program performance measures. In 1996, the Head Start Bureau established a Performance Measures Center which is using these instruments to conduct FACES.

FACES, a longitudinal study of 3200 children and families in 40 nationally representative Head Start programs across the country, launched in the fall of 1997, will provide comparisons of Head Start children's developmental status with their same-aged peers at program entry, exit, and after 1 year of kindergarten, in the following ways: a) comparison with overall nationally normed data on tests of vocabulary, prereading, math skills, and prewriting; b) comparison with children from low-income families who have not attended Head Start based on parent reports of accomplishments and difficulties (National Household Education Survey, 1993, 1996) and on published mean scores from direct child assessments of low-income samples; c) comparisons among Head Start children who participate in the program to different degrees or for varying lengths of time, for example, 1 versus 2 years; d) comparison of the rate of development shown by Head Start children and all preschoolers of the same ages.

In addition, the use of observational measures of program quality, often used in national studies of preschools and child care centers, will provide Head Start with ongoing assessments of the educational quality of classroom environments. Finally, the opportunity to link indicators of program quality with child outcome data over time will provide the strongest test of program accountability.

Head Start is partnering with the Department of Education's Early Childhood Longitudinal Study (ECLS-K) which will draw a national sample of 23,000 kindergarten children in the fall of 1998,

4

including 3000 former Head Start children and a comparable group of eligible nonattendees. Using the same assessment battery for ECLS-K and the FACES kindergarten follow-up, we will be able to compare child outcomes in kindergarten and the early school years between children in Head Start and children who attended other child care or preschool programs, or who had no preschool experience.

Finally, Head Start is undertaking a set of locally-based feasibility studies for determining Head Start effectiveness through its QRC. These feasibility studies were outlined in a response to a recent GAO report. These impact initiatives are presented in more detail starting on page 11 of this response.

The Head Start research agenda has been shaped over the years based on input from a number of leading researchers from the field. The various researchers that have been consulted consistently have recommended against conducting a single nationally-representative impact study of the effectiveness of Head Start due to a variety of ethical, methodological and logistical difficulties. In fact, a host of methodological threats to the assumptions of a random assignment design, outlined below, may make the outcomes of such a study less reliable or valid than the currently proposed research design.

First, recruitment of children and families for Head Start is guided by mandates to serve the most needy of eligible families and to incorporate a minimum of 10 percent of children with disabilities in each program--half of whom would be recruited, but unserved in a random assignment design. Second, in some States and many communities, Head Start is already serving more than a majority of eligible children and families. In these situations, it would be impossible to conduct the form of evaluation recommended by GAO because of a lack of sufficient numbers of unserved children to comprise a control group. Third, the mobility of low-income populations is likely to result in children leaving treatment in unpredictable numbers and at unpredictable time points, or crossing over between program and comparison groups. Finally, randomization also assumes that everyone in the treatment group and everyone in the control group gets the same treatment (program) within their group. However, based on available data, it is virtually assured that control group children will participate in a wide range of other early childhood programs and settings, of varying levels of quality, intensity, and comparability to the standard of Head Start services.

In 1990, the Advisory Panel for the Head Start Evaluation Design Project put forth a set of research recommendations in the publication "Head Start Research and Evaluation: A Blueprint for the Future."

5

> *"An overall research strategy rather than a single large scale study is the appropriate framework for addressing critical Head Start research and evaluation questions. The Panel recommends strongly against a single large-scale study of Head Start as the principal mechanism for seeking answers to the pivotal research questions highlighted above. The methodological requirements for the new generation of research and evaluation issues do not lend themselves to large-scale evaluations that treat Head Start as a single program. Head Start is not, in any simple sense, a uniform "treatment" . . . .*
>
> *It is important to note then, in the view of the Panel, randomized studies designed to compare the effects of Head Start against the effects of nonparticipation ("treatment vs. no treatment") are generally no longer viable options. First, as ACYF progresses toward the Administration's goal of universal services for all eligible children, the potential for withholding services to form a control group, already difficult for ethical and practical reasons, will cease to be an option. Second, in view of the expansion of State and public school preschool programs and developmental child care, it is unrealistic to expect to find in most communities a representative group of "untreated" eligible children, even if Head Start services are not provided.*

The current body of research on early childhood interventions does achieve consensus on some points. When studying program effects, it is imperative to take into account such factors as family and neighborhood characteristics, culture and language, quality and intensity of services received (including observation of classrooms and caregiver-child interaction), as well as parent participation in and satisfaction with the program. There is less consensus about the best or most feasible methodology for answering questions about program effectiveness, especially given the various programmatic, methodological and/or logistical threats to the validity of the data derived from many of the available design options. Some researchers advocate using comparisons to nationally normed instruments or large-scale studies, while others insist that random assignment designs are the standard, however difficult to carry out in real-life situations. Given this divergence of opinion on such an important question, perhaps the best recourse would be to seek consensus and advice from experts called together by an external scientific organization, such as the National Academy of Sciences.

6

Strategies for Head Start Monitoring

Section 641A(c) of the Head Start Act contains a requirement that each Head Start grantee receive a full on-site review of all program and fiscal requirements at least once every 3 years; therefore, approximately one-third of all grantees are monitored each fiscal year. In addition, new Head Start grantees are reviewed immediately after completing 1 year of providing Head Start services.

The purpose of monitoring is to assess each grantee's compliance with Head Start regulations and to strengthen the overall quality of Head Start services provided by each grantee. On-site monitoring reviews are a key element of Head Start's overall strategy to ensure quality services and local accountability.

We believe that we have a strong and effective strategy for carrying out our legislative requirements for monitoring Head Start grantees. A Head Start on-site monitoring review examines a grantee's ability to provide quality services consistent with current Head Start regulations. Using a team approach and a standard monitoring instrument, program operations are reviewed through a process of observations, record reviews and extensive interviews with a variety of individuals including program managers and other staff, parents and other members of the community in which the program operates.

Through these processes, the system that is used for Head Start monitoring accomplishes what it is intended to do. It is comprehensive in scope and contains a variety of checks and balances to minimize inappropriate subjectivity in review findings. It is effective in identifying problem grantees and in setting the stage for a course of action to correct problems. It allows a grantee opportunities to correct problems that have been identified before adverse action is taken to terminate them and, through the training and technical assistance network, supports a grantee in making changes that are necessary.

Building on what works well in the current monitoring system, the Department is taking steps to improve and strengthen the monitoring process and the system that supports this effort.

Revisions to the Head Start Program Performance Standards, the core set of regulations for all Head Start programs, created an opportunity to review and rethink the process that is used for monitoring, including the instrument that is used. Since the revised Performance Standards went into effect on January 1, 1998, grantees have been reviewed using an interim instrument that follows the revised standards. While this is occurring, the Head Start Bureau is developing a revised instrument and process that will be used to monitor grantees in the future.

7

We have learned, largely from working with grantees with deficiencies, that addressing systemic issues and their impact on services has a greater impact on the overall quality of a program than identifying separate and very individual and nonrelated, noncompliant items. Because of this, the revised process will take a "systems approach" to reviewing Head Start grantees. The systemic approach will first look at how services are organized to implement the Performance Standards and other regulations. At the same time, review teams will be looking at how the Head Start organization's management systems impact services. From working with the concept of deficiencies, we know that many issues that are identified in program services are often caused by problems within the management structure of the organization. For example, incomplete health services to children could be caused by a lack of program planning to assure that services occur timely. It could also be related to poor or incomplete record keeping or poorly trained staff. We believe that this approach will better enable grantees to address and correct problem areas and will have a greater impact on quality within all Head Start programs.

We are also working to strengthen the training that reviewers receive. This will help ensure that all reviewers understand the process and have the knowledge and skills that are necessary to function as a review team member.

Current Head Start Impact Initiatives

Although GAO's report provides a brief overview of the ways in which Head Start's current and proposed research initiatives will answer the question of Head Start effects on child outcomes, the Department would like to expand and clarify the range of its approaches to this question. Taken together, we are convinced that this approach will provide rich and definitive data to answer the question of whether Head Start is making a difference in the lives of the children and families it serves.

Family and Child Experiences Survey

FACES will be able to compare Head Start children's developmental status with their same-aged peers in the following ways:

a. *comparison with overall age-sex norms:* these are comparisons between the assessment scores of Head Start boys and girls of specific ages (3, 4, or 5) with the average scores of all boys or girls of the same age. Such comparisons will be possible for scores on the Peabody Picture Vocabulary Test, the Woodcock-Johnson sub-tests, the McCarthy subscales, and the Test of Language Development phonemic analysis subscale, and the Social Skills Rating Scales;

8

b. *comparison with children from low-income families who have not attended Head Start:* the availability of national survey data on both the developmental status and program attendance histories of subgroups of preschool children make possible comparisons between Head Start children and those from low-income families who have: a) attended no center-based preschool program; or, b) attended another type of preschool program. Such comparisons will be possible for parent reports of children's developmental accomplishments and difficulties, which were collected in the National Household Education Survey (including both 1993 and 1996 data), in addition to comparisons with published data on mean scores from direct assessments of low-income studies;

c. *comparisons among Head Start children who participate in the program to different degrees or for varying durations:* it is also possible to use the FACES data themselves to develop "dose-response curves." These involve comparisons of the developmental progress of children of the same age who attend Head Start classes for relatively few hours per week with those who attend for a greater number of hours each week. Similarly, the development of children of the same age who have attended Head Start for 2 years can be compared with the progress of those who have attended for only 1 year; and,

d. *comparisons of the rate of development shown by Head Start children and all preschoolers of the same ages:* as longitudinal data are collected in FACES, it will be possible to use national normative information from the child assessment measures to compare not only the level of development, but also the slope or rate of development, displayed by Head Start children and national samples of similar ages and sexes. For example, even though some Head Start children may be behind their agemates in the level of their development, they may show equivalent amounts of growth between the fall and spring of the Head Start year, or from one year to the next.

In addition to the comparisons enumerated above, a major thrust of the FACES analysis will be to relate differences in children's development and family behavior change to program quality measures and other aspects of the Head Start centers and programs the children attended.

9

Early Childhood Longitudinal Study-Kindergarten and Birth (ECLS-B)

Head Start is also partnering with the Department of Education's Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K), which will draw a national sample of 23,000 kindergarten children in the fall of 1998. Head Start has funded a verification study within ECLS-K, to assure that children identified by their parents as having previously attended Head Start had actually been enrolled in Head Start preschool programs. ECLS-K anticipates that 3000 children in the study will have been enrolled in Head Start, with a comparable number of children having been eligible for Head Start, but having attended either no preschool or another preschool or child care setting. Using a similar instrument battery for ECLS-K and the FACES kindergarten follow-up in spring 1999 will provide useful comparisons in an extremely timely manner.

It will be possible with the ECLS-K data to compare the school progress of children who previously attended Head Start with that of children from low-income families who had no center-based program experience. While ECLS-K is not a study of Head Start programs, this comparison should help elucidate some of the longer-term effects of Head Start participation.

Both FACES and ECLS-K concentrate on assessing children in the same developmental domains, which have been indicated as key by both the Department of Education's Goals 2000 and the Head Start Performance Measures: 1) Language/Literacy Dimension, including receptive vocabulary, prereading, letter-sound relationships, print awareness, prewriting; 2) Quantitative Ability/Mathematical Thinking Dimension, including knowledge of numbers and counting, numerical memory, arithmetic calculations; 3) General Knowledge Dimension, including basic facts about self and family; 4) Social Skills Dimension, including cooperation, assertion, responsibility, empathy, self-control, self-esteem, social problem solving; 5) Problem Behavior Dimension, including difficulties with attention, aggression, anxiety, low-self-esteem, and impulsivity; 6) Fine and Gross Motor Development Dimension, including eye/hand coordination and perceptual-motor skills; 7) Health Dimension, including general health, chronic illnesses, nutrition and physical activity, and 8) Physical Development Dimension, including height, weight, vision, and hearing.

Head Start is also working with the National Center for Education Statistics (NCES) as part of an interagency workgroup planning the ECLS-B, with a national birth cohort to be drawn in 2000. Head Start has already contributed to the study by offering an extensive battery of measures developed over the past 2 years through its Early Head Start Research and Evaluation Project

national evaluation contract, an evaluation of the Early Head Start program serving infants and toddlers. Dialogues with NCES are ongoing about such issues as research sample, design, frequency of data collection, and intensity of measures used.

Quality Research Center Consortium Effectiveness Studies

The Head Start QRC, consisting of four Quality Research Centers (QRCs) at 1) the Family and Child Care Research Program, Frank Porter Graham Child Development Center, University of North Carolina, Chapel Hill; 2) High/Scope Educational Research Foundation in Ypsilanti, Michigan; 3) Education Development Center, Inc. in Newton, Mass.(a consortium including Harvard University, Boston College, and the Massachusetts Society for Prevention of Cruelty to Children); and 4) Georgia State University, is currently in the planning stages of designing and conducting a set of locally-based feasibility studies for determining Head Start effectiveness. These feasibility studies were outlined in a response to a recent GAO report. Having already established partnerships with local Head Start programs as part of the consortium's focus on quality research over the past 2 years, the QRC sites are currently in the process of working with program partners on the feasibility of implementing effectiveness studies utilizing a random assignment design. To date, the QRCs have focused their efforts on such areas as gaining permission from Head Start policy councils in each program, investigating current program recruitment and enrollment practices (including criteria for program acceptance), designing alternative lottery-type enrollment procedures, identifying other preschool and early childhood programs available in various communities, (and the percent and characteristics of children and families enrolled in such programs), and deciding on assessment protocols which build on the previous work of the QRCs, FACES, and links to other national studies. While there may be common elements across these locally-based studies, there will also be variability based on the need to respond to conditions specific to each community, and to questions of particular local interest.

Appendix II

GAO Contacts and Staff Acknowledgments

GAO Contacts

Harriet C. Ganson, Assistant Director, (202) 512-9045
Veronica Scott, Evaluator-in-Charge, (202) 512-7039

Staff Acknowledgments

In addition to those named above, the following individuals made important contributions to this report: Karen A. Whiten, Assistant Director; Mary Roy, Senior Evaluator; Donnesha Y. Correll, Issue Area Assistant; C. Jeff Appel, Adviser; and Liz Williams, editor.

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. VISA and MasterCard credit cards are accepted, also. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 37050
Washington, DC 20013

or visit:

Room 1100
700 4th St., NW (corner of 4th and G Sts., NW.)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (202) 512-6061, or TDD (202) 512-2537.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touch-tone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

or visit GAO's World Wide Web Home Page at:

http://www.gao.gov

BEST COPY AVAILABLE

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested

BEST COPY AVAILABLE